DOCUMENT RESUME

ED 109 339 CE 004 118

AUTHOR Krueck, Thomas G.
TITLE A Comparison of Three Methods of Conducting a Follow-up Study.
PUB DATE Apr 75
NOTE 34p.; Paper presented at the annual meeting of the American Educational Research Association (Washington, D. C., April, 1975)

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS *Comparative Analysis; Cost Effectiveness; Data Analysis; *Data Collection; *Educational Research; *Followup Studies; *Graduate Surveys; Interviews; Questionnaires
IDENTIFIERS *Telephone Surveys

ABSTRACT

With the increased use of the follow-up study on graduates, it has become necessary to research and compare the results of alternative methods of data collection on a diverse student population. The purpose of this study is to determine the most effective data collection method by examining the rate and nature of responses and the cost of three methods--personal interview, mailed questionnaire, and telephone survey. The study disagrees with previous research indicating that the mailed questionnaire is the most appropriate method. A comparative analysis of the three methods of conducting a follow-up study of former high school students at the Skyline Career Development Center (CDC) was designed to compile and design survey instruments, to incorporate and document the appropriate methodologies, and to obtain meaningful information on graduates and dropouts to provide a basis for program alteration. From the evaluation of cost, data analysis, findings, and interpretation, it was found that telephone surveys combined with mailed questionnaires elicited a higher and more critical response rate at less cost for both small and large districts. For best results, recommended basic procedures are: accurate identification of the student population, careful use of clerical personnel, and goal-oriented instrumentation. (JB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

APR 21 1975

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A COMPARISON OF THREE METHODS OF CONDUCTING A FOLLOW-UP STUDY

Thomas G. Krueck

Dallas Independent School District

ED109339

The practice of conducting a follow-up survey of graduates is neither new nor innovative. Local school districts, governmental agencies, and independent researchers have been involved in such studies for over four decades. Recently, there has been a renewal of interest in the use of follow-up surveys. The increasing emphasis on accountability imposed on educators by legislators and the general public may be an important impetus for more research and evaluation activities in which a follow-up survey is used. Follow-up studies become an after-the-fact needs assessment based on the evaluation of existing programs by former students. Data of this type can only be collected through a follow-up survey. Such data are indispensible if one assumes that former clients are worthy of consideration in such matters. Therefore, the utility of a follow-up survey can be dichotomized into (a) responding to pressures for accountability with factual data and (b) fulfilling the practitioner's constant need for data through which he can implement program improvements. In fact, the State of Texas now requires a five-year follow-up on all graduates of vocational programs funded by the State.

With the understanding that such studies have a place in research and evaluation, it is unfortunate that follow-up studies are typically conceived as a cursory examination of what graduates are doing one to five years after

---

Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C., April 1975.

CE004118

leaving school. Such studies are burdensome, costly, and frequently are severely limited by poor response rates. While a great deal of research has been conducted about the methodology and costs involved in follow-up studies, little, if any, of this knowledge has made its way to local school districts. Hence, application lags far behind research.

It would appear that the educational community needs research about follow-up techniques that is generalizable and relevant to its setting. This need can best be met by comparing the results of alternative methods of data collection on a diverse student population. An examination of rate of response, cost, and nature of response would allow school districts (regardless of size or student composition) to select the method most appropriate for their purposes.

One other aspect--intent--must be discussed. Historically, follow-up has involved only counting heads and identifying the activities in which former students are presently engaged. It is the belief of this researcher that follow-up studies should provide meaningful data that will allow for program alteration; consequently, the present study was designed to accurately compare data obtained using three well-known methods of collecting follow-up information. Each method employed--personal interview, mailed questionnaire, and telephone survey--utilized existing knowledge concerning the most advantageous way to obtain a high response rate. An examination of the results and instruments should prove useful to educators planning any type of follow-up survey.

The present study is an attempt to satisfy the unmet need so aptly cited by Jackson and Rothney (1961) in the following statement:

> The merits, weaknesses, and biases of both the mailed questionnaire and the interview have been discussed frequently, but there has been little research designed to assess the comparative contributions of the two procedures for follow-up purposes.[1]

While the above comment was made in 1961, it is just as timely thirteen years later. In fact, it could be expanded to include the telephone survey which has been the subject of even less research.

While there are exceptions, follow-up studies generally continue to be conducted on a large scale with disappointingly low percentages of return. Consider the following sample of studies:

1. Four mail surveys of East Bakersfield High School graduates of 1947 and 1948 yielded returns ranging from 25 to 39 per cent.[2]

2. A state-wide study of 14,000 Utah high school graduates utilizing advertising to increase returns obtained a return of 52 per cent of the questionnaires.[3]

3. A study of the Highline District graduates of 1970 yielded only a 68 per cent response after replacement of unattainable graduates.[4]

4. A study of the Syracuse seniors shortly before graduation in June, 1970, has a return rate of only 45.7 per cent.[5]

5. Only 49 per cent of a random sample of Tacoma households responded to a survey conducted by the local district.[6]

The returns for some of the studies cited above may seem relatively good; however, extenuating circumstances exist in those instances. For example, the study in Syracuse was conducted while students were still in school. The Highline study continually substituted respondents for those who were unattainable and still only obtained responses from 68 percent.

Several points become apparent when one considers the research reported

in the area of follow-up surveys. First, little is known, or at least little has been reported, about the use of the telephone in follow-up surveys. Second, the information presented in a prior review of research indicated that, for large-scale surveys, the mailed questionnaire is presently considered the most appropriate. Finally, only a small portion of the research in the area of follow-up surveys has dealt with comparisons of methodology and technique; and those studies appear to have serious limitations for utilization by practitioners in local school districts.

Statement of the Problem

Based on work in and with local school districts, certain assumptions were made concerning the present situation with regard to follow-up research. These were:

1. Some form of follow-up study is conducted in most sizable school districts. The reason for such studies may vary greatly from district to district.

2. Monetary and personnel constraints often limit all but the largest districts in the conduct of sophisticated follow-up research. However, this does not lessen the need for the research in smaller districts.

3. Disappointing results are rarely published or released to outsiders. If one accepts this assumption in the light of the results previously enumerated, this accentuates the problem that presently exists.

The problem is that follow-up research--a useful and accepted tool--is often abused or misused by those who could benefit greatly from such research. Therefore, the intent of this study is not to devise new or innovative procedures for carrying out a follow-up survey but, rather, the intent is to identify the most efficient follow-up survey method, or combination of methods,

based on cost, rate of return, and nature of response. Instrumentation and selection procedures for each method examined are, of course, concomitant products of any such study. However, the procedures and instruments represent only an application of the existing body of knowledge available to any researcher.

For nearly four decades, follow-up studies have been conducted by local school districts and yet few, if any, changes have taken place regarding methodology. Typically, this methodology takes the form of sending questionnaires out via bulk mail and waiting for their return. After patience falters under the pressure from administration, results are released--often with the omission of the percentage response or the assurance that the one-third returned are a fair representation of the entire population. As early as 1942, Nichols responded to the situation thusly:

> Why set great store by the results of a follow-up study that got returns from but 59.9 per cent of the group studied? Why not take cognizance of the fact that those who are least proud of their after-school adjustments are least likely to reply?[7]

While a 100 per cent response rate seems unlikely, it is only reasonable to strive for results in excess of 60 percent since school administrators continue to make decisions based on returns as low as 30 per cent and rarely greater than 50 per cent.

It was the purpose of this study to seek answers to the following questions:

1. Which method of conducting a follow-up study elicits the greatest percentage of response?

2. Are there any differences in the nature of responses elicited by each method of conducting a follow-up study?

3. What are the comparative costs for each method?

4. Which method or combination of methods is most cost-effective for conducting a large-scale follow-up study?

5. Are some methods more effective than others for certain ethnic groups?

6. What would be the most advantageous design for a typical follow-up study?

In addition, it is hoped that the versatile set of survey instruments developed will have application potential, with minor alterations, for any secondary or vocational school follow-up survey of graduates.

It is hoped that the above-mentioned contributions will encourage more reliable studies at a known cost so that decisions can be based on the best possible data. Another need for this study is found in the continually increasing demands being placed upon local districts to provide follow-up data concerning graduates and drop-outs. As mentioned earlier, the State of Texas requires a follow-up study conducted over five years on all students who were enrolled in "approved" vocational training courses. However, no detailed plans for conducting the follow-up have been provided since the issuance of guidelines by many states during the 1940s and early 1950s. There also exists some question as to what information is most desirable. There presently is a preoccupation with what former students are doing rather than the inclusion of questions concerning what they might recommend in the area of program development.

Limitations of the Study

The most obvious limitation apparent in this study, as well as other similar studies, is the inability to assure that inroads will be made into

existing practices--even though the research is designed to produce results useful to public school districts. Another limitation is that each method may be approached differently and therefore the findings reported here will only be applicable to studies similarly conducted. However, steps have been taken to assure that the best-known procedures will be employed for each method.

Finally, a limitation previously alluded to pertains to the instruments. Each instrument, while alterable, was designed for use with the specific population under study However, it was not the intent of this study to provide a set of questions applicable, without changes, to any and all secondary school settings.

Design of the Study

In review of the limitations of the study discussed in the preceding section, it was imperative that specific goals and objectives be established for the study. The design of the study would then follow as a logical extension of these goals and objectives, incorporating proven techniques to ensure the utility and generalizability of the findings.

The following goals and objectives were specified as the basis for and the intent of the study:

Goal I. To conduct a comparative analysis of three different methods of collecting follow-up data from former high school students.

Objective 1: Identify the response rate for each of the methods under study.

Objective 2: Assess the advantages or disadvantages of each method in terms of the actual responses obtained.

Objective 3: Provide a cost-effectiveness measure for each of the methods employed.

Goal II. To compile and/or design instruments which, with minor alterations, could be used as a format in conducting a student follow-up program.

Objective 1: Identify the various types of information to be obtained from respondents.

Objective 2: Establish a suitable mail questionnaire which could, with minor revisions, be made applicable to other programs.

Objective 3: Establish a suitable interview schedule which could, with minor revisions, be made applicable to other programs.

Objective 4: Establish a suitable telephone survey which could, with minor revisions, be made applicable to other programs.

Goal III. To incorporate and document the methodologies most appropriate for carrying out each survey technique.

Objective 1: Extract from previous studies the most advantageous method for conducting a mail questionnaire survey.

Objective 2: Extract from previous studies the most advantageous method for conducting a personal interview survey.

Objective 3: Extract from previous studies the most advantageous method for conducting a telephone survey.

Objective 4: Document the actual process of data collection required for each survey technique.

Goal IV: To obtain meaningful information from graduates and dropouts from the Skyline Career Development Center that would provide a basis for program alteration or, at least, lend impetus to considering such changes (see the sample section for clarification).

All of these goals were part of a more ambitions, and hopefully achievable, goal--to increase the number of sophisticated and appropriately conducted follow-up studies of students by local school districts.

Sample

The Skyline Career Development Center (CDC), a school-based career education program in Dallas, Texas, was selected for the study because of the availability of needed data and because the student body included a large number of minority students.

All graduates and seniors who did not return to the CDC after the 1971-72 school year were included in the study. The rationale for including all such persons was that the opinions held and activities engaged in by any student are worthwhile and meaningful. Furthermore, to ignore an individual simply because he failed to graduate would be a serious criticism of any follow-up study.

After preliminary examination of this group, a total of 492 former students were identified who had completed at least one year of study in the CDC and who were no longer enrolled in the secondary schools of Dallas. Nine potential subjects were dropped from the sample because of such reasons as death, long-term illness, and mental disorders. Table 1 shows the sex and ethnic background of the sample.

All persons included in the population under study were sent a preliminary questionnaire and a letter of explanation concerning the study in November, 1972. Approximately 40 per cent of the persons failed to respond; however, data were obtained on all but 27 of the subjects through relatives and friends.

The inability to contact subjects via mail was due primarily to the following reasons: students had left home, their families had moved, or information in the school's file was inaccurate as to the student address and/or telephone number. The 27 non-respondents were kept in the study on the basis that they might be reached in the more structured treatments carried out over a longer period of time.

Table 1

Classification of Post-High School Population Under Study by Ethnic Background and Sex

	Sex					
	Male		Female		Total	
Ethnic Background	n	Percent of Total	n	Percent of Total	n	Percent of Total
Anglo	171	47.9	186	52.1	357	72.6
Black	23	30.3	53	69.7	76	15.5
Mexican-American	21	36.8	36	63.2	57	11.6
Other	1	50.0	1	50.0	2	0.3
Total	216	43.9	276	56.1	492	100.0

Assignment to Treatments

The three methods of data collection under study (hereafter referred to as treatments) and the major reason for their selection were as follows:

1. Mailed Questionnaire--selected because it is the most widely researched and utilized method of conducting follow-up research.

2. Personal Interview--selected because it should allow for maximum contact between researcher and respondent.

3. Telephone Interview--selected because it presumably combines desirable features of both the mailed questionnaire and personal interview techniques.

The 492 subjects selected for the study were randomly assigned to the three treatment groups. To ensure against biasing by assignment, no preferential assignments were made based on prior knowledge obtained from the data collected in November, 1972. Table 3 identifies the composition of each treatment group based on sex and ethnic background. The random assignment did place a slightly higher than proportionate number of Blacks in the mailed questionnaire treatment and Mexican-Americans in the personal interview treatment group.

To examine whether or not substantive differences would occur based on treatment received, 48 subjects (16 from each treatment) were randomly selected to receive a second treatment. This selection was made before any data collection had begun. After the completion of data collection for all groups, attempts were made to obtain information utilizing the second treatment. These attempts were made even if no data were obtained as a result of the first treatment. By collecting the data in this fashion, the order of treatment was counterbalanced, thereby eliminating the possibility of biasing the results.

Instrumentation

In an effort to ensure that the data obtained from each treatment would be comparable, the content of the instruments for each treatment was kept equivalent. Only minor changes necessitated by the type of treatment were made.

Content. The data obtained from the instruments covered each of the following:

1. Background Data

    a. Birthdate
    b. Sex
    c. Ethnic Background
    d. Area of High School Specialization
    e. Social Security Number

2. Possible Status Changes

    a. Address
    b. Phone Number
    c. Marital Status

3. Present Employment Status

    a. Where Employed
    b. What Type of Work
    c. Duration of Employment

4. The Perception of Former Students with Regard to CDC

    a. Strengths
    b. Areas in Need of Improvement
    c. General Level of Satisfaction

5. Future Plans

    a. For Personal Advancement
    b. Mobility
    c. Notification of Future Study

Pilot Test. To ascertain the utility of the newly developed instrumentation, a pilot study was conducted. A sample of thirty twelfth-grade Skyline Center students (ten from each ethnic group) were selected across various career clusters. In the spring of 1973, interviews were obtained from these students utilizing the instrumentation.

The following changes were indicated as a result of the findings of the pilot study:

1. The five-point scale for assessing firmness of future plans was changed to a four-point scale to eliminate clustering.

2. Response categories were expanded to cover alternative options.

3. Some questions, such as length of time on the job, were made closed-ended for ease of obtaining and coding responses.

4. Alternative terminology was developed where it was found that terms such as "greatest strength (asset)," and "ambitions," did not always communicate with the respondents.

Some questions were eliminated because they were too time-consuming and because the information obtained was of questionable value for decision makers.

Refinement. Items on all forms of the instrument included both open- and closed-ended questions. The choice of each type was determined by considering the response time required and the nature of possible responses to the item. Questionnaire items were presented in an open-ended format where responses were potentially quite varied or where respondent wording was desirable. Closed-ended questions were incorporated primarily for three reasons: (1) they require less time to administer, (2) they are generally less formidable to respondents than questions requiring an extended reply, and (3) closed questions tend to eliminate coding difficulties.

All questions were constructed and pilot tested to minimize the possibility of making the following item errors:

1. Failure to word questions in a manner which would convey an accepted meaning.

2. Wording beyond the respondents' ability to comprehend.

3. "Leading questions" that would elicit responses that were desired on the part of the researcher.

4. Items that combined two or more questions in a single item.

5. Questions that might lead the respondent to either refuse to respond or prevaricate in his or her response.

After the above changes had been incorporated, the proposed final forms of instrumentation were reviewed by a group of Skyline seniors who felt that the instruments did, in fact, communicate; i.e., were clear and unambiguous.

Methodology

Since assignment to treatment occurred independent of such considerations as whether a respondent could be reached by telephone or was located a great distance from Dallas, it was anticipated that a number of potential respondents would be missed. However, random assignment to treatments was necessary to ensure that all treatments could be compared without biasing outcomes. Future follow-up studies might wish to take such considerations into account and make appropriate alterations to improve response rates.

The interviews, telephone calls, and mailings took place during the months of May through August, 1973. To ensure that each individual received similar treatment, only four interviewers, each with previous interview experience, were used. The interviewers constantly interacted with each other to analyze and recommend methods to ensure the highest quality of response and consistency in interviewer technique. A detailed description of the procedures employed for each treatment method follows.

Analysis of the Data

Once the data were collected for all three treatment groups, typical data preparation procedures were employed. The first step required coding of the information so that computer applications could be utilized to categorize and analyze the data.

To minimize the possibility of differences in interpretation of responses, all coding was done by two individuals. Higher reliability was anticipated by employing only two coders. To ensure validity in coding, a sample of completed instruments were submitted to third parties for coding--the intent being to see if a consensus existed as to the appropriate coding of open-ended questions. This precaution, however, did not replace the normal series of random checks to ensure interrater reliability.

All coded data were then keypunched and verified for computer use. After initial computer runs, which included simple cross-tabulations, it was necessary to correct some erroneous coding and to adjust for the large number of responses coded as "other" due to the channeling of responses to open-ended questions into eight possible response categories. While some specificity was lost at this stage, the gains in utilization of findings warranted limiting the possible response categories.

The computer analysis of the data involved the establishment of a series of cross-tabulations based on sex, ethnicity, curriculum background, and treatment. This was done to prepare the responses for data analysis utilizing the chi-square test of goodness of fit for independent samples. The chi-square test was chosen as the most suitable test for analyzing the data since the data were in frequencies and the measurement of the variables included those in a nominal scale and in discrete categories of an ordinal

scale. All other tests commonly used for tests of independent samples require at least ordinal measurement of the variables; hence, the chi-square test is uniquely useful for data such as that collected for this research.

In essence, the chi-square test was used to ascertain if the samples had come from the same population or from identical populations with respect to the proportion of cases in the various categories involving response rates for the various treatments, effectiveness of these treatments with different ethnic groups, and nature of responses obtained.

Questions concerning cost and cost-effectiveness were examined by totaling costs and dividing by the number of responses. Finally, the analysis of the most advantageous design for future studies involved the weighing of information obtained in examining previous questions.

Findings

Question One

The first question addressed was: Which method of conducting a follow-up study elicited the greatest percentage of response?

Attempts were made to follow-up a total of 492 former students of the Skyline Career Development Center. After pursuing the methods outlined previously, 269 individuals finally responded to the survey. This gave an overall response rate of 54.7%. The null hypothesis that the proportion of responses was the same for the three modes of follow-up was tested, and a statistically significant difference ($\chi^2 = 12.71$, $df = 2$, $p < .005$) was found among treatments. Examination of the data indicated that the mailed questionnaire and personal interview approaches were approximatley equal in effectiveness eliciting response rates of about 50 percent. The telephone

survey method was considerably more effective than the other two methods with a response rate of approximately 66 per cent. Further analyses of the data comparing the frequencies for each possible pair of treatments indicated that the telephone survey technique was significantly more effective than either the mailed questionnaire or personal interview methods, while no significant difference was found between the latter two treatments.

The reasons for not receiving responses from subjects varied among methods. The most important reasons are listed in Table 2 with the number of non-respondents for each reason shown by treatment. Obviously, these reasons are not necessarily mutually exclusive, since some potential respondents who were never contacted might also have ended up in the "refuse to respond" category.

Table 2

Comparison of Occurrence Among Treatments with Respect to Reasons for not Receiving Responses

Reason	Mailed Questionnaire	Personal Interview	Telephone Survey	Total
No telephone	1	0	3	4
No forwarding information	0	11	3	14
Wrong number or disconnected telephone	12	5	11	28
Out of Dallas area	5	13	19	37
Could not follow up	24	30	14	68
Refused to respond	3	8	1	12
Total number of non-respondents	81	86	56	223

Inspection of the data displayed in Table 2 reveals that methods requiring some form of personal contact with the subject were less effective when dealing with subjects who had left the immediate area; however, this was to be expected since no long-distance calls were. Other interesting findings showed that:

1. There were appreciably fewer respondents who could not be followed up via the telephone as opposed to the other two methods.

2. Respondents were most likely to refuse to respond when asked for a personal interview.

3. Quality of data (i.e., phone numbers and addresses), inability to follow up (i.e., line busy, no answer, etc.), and respondents living outside the Dallas area were the major reasons for non-response.

Question Two

The second question posed was: Were there any differences in the nature of responses elicited by each method of conducting a follow-up study?

This question was posed to examine two aspects of responses. The first dealt with communication; that is, the extent to which respondents gave no response to various questions. The second aspect examined whether or not personal contact, either by telephone or in person, would inhibit respondents, thereby lessening the chance that a respondent would comment negatively on certain subjects. The randomization of subjects to treatment eliminated potential bias in favor of any one method of seeking response, thereby allowing for examination of differences in the nature of response.

The extent to which respondents were able to formulate answers to questions, as opposed to failing to respond, varied significantly from one

Treatment to another. Answers to questions concerning satisfaction with instruction and counseling, and open-ended questions dealing with the Skyline Center, showed that former students responding to a mailed questionnaire had a much higher frequency of no response, especially for questions concerning the Center.

Two questions were identified as suitable for assessing the willingness of respondents to make negative comments. These questions involved the rating of instruction and counseling on a five-part scale ranging from excellent to poor.

There were significant differences in the number of negative responses received depending on treatment and the question being asked. Evaluation of instruments showed that the only statistically significant difference ($\chi^2 = 4.83$, df = 1, $p < .05$) on a paired comparison basis existed between the mailed questionnaire and personal interview approaches. No statistical significant results were obtained when comparing the mailed questionnaire against the combined efforts of the other two methods. The rating of counseling differed in that significance ($\chi^2 = 11.48$, df = 1, $p < .005$) was found between the mailed questionnaire and the telephone interview. However, the mailed questionnaire had a statistically significant difference in the number of negative responses when compared to the combination of personal forms of data collection. Therefore, while the mailed questionnaire did elicit a greater number of negative responses than personal forms of data collection, the lack of any consistency across the two questions, particularly when comparing the mailed questionnaire and the telephone survey results, indicates that further research is needed with ragard to the effect of personal involvement on willingness to respond in a negative mode.

Question Three

The third question dealt with costs: What were the comparative costs for each method?

Fixed Costs

In examining this question, it is important to recognize that there were fixed costs, such as printing and processing, that applied commonly to all three treatments. While these costs did not vary among treatments, their existence must be considered in cost estimates. The fixed costs, however, were relatively small since the District provided printing and computer time to the investigator at cost. Total costs for printing and processing were approximately 20 dollars for printing and eight dollars for computer time.

Manhour costs can be broken down into three categories: (1) prior to treatment, (2) during treatment, and (3) after treatment. Each category will be examined separately because time can be saved in future implementations since certain developmental stages used need not be repeated.

Furthermore, the types of manpower required varied greatly depending on the task at hand. For ease of providing meaningful comparisons of manpower costs, clerical level costs were calculated at three dollars per hour, research assistant costs were fixed at five dollars per hour, and senior researcher costs were set at ten dollars per hour.

Prior to treatment, the research staff was required to construct an instrument, randomly assign former students to treatments, and plan a design. Approximately 60 hours of work (including pilot testing) were required. This included 50 hours of senior researcher time, five hours of research assistant work, and five hours of clerical help. In the future,

this stage should require less than eight hours since instrumentation has been completed.

During treatment, monitoring of staff activities required a great deal of time on the part of the investigator. It is impossible to estimate precisely the total number of hours spent in this activity. However, the total amount of time was extensive, involving a portion of each day during the time data were being collected. This was considered as a role responsibility of the senior researcher involved in the study.

The greatest amount of required manpower was needed after the collection of data was completed. This entailed the coding, keypunching, verifying, and analysis of the data collected. The time required for these tasks were:

1. approximately 57 hours of clerical assistance for coding,
2. approximately 10 hours of clerical assistance for keypunching,
3. approximately 40 hours of clerical assistance for verification, and
4. approximately 60 hours of senior researcher time for analysis.

Verification and analysis of the data were greatly facilitated through the use of electronic data processing. It would be safe to assume that if a computer were not available, the time required for manipulation of data by hand would inhibit much of the analysis necessary for meaningful interpretation.

Treatment Costs

Costs peculiar to treatment varied greatly, depending on the treatment involved and particularly the methodology employed.

The initial response to the mailed questionnaire was 37.8 per cent. After one telephone reminder, the response rate increased to 42.7 per cent. A second telephone reminder brought an additional 7.3 per cent response.

A third call was made to one student who subsequently sent his questionnaire back. This process obtained 83 responses out of a possible 164, or a 50.6 percent response rate.

A total of 164 questionnaires were sent out initially at a cost of 16 cents each (this cost included the self-addressed, stamped return envelopes). Seventeen additional questionnaires, of which nine were returned completed, were sent upon request at a cost of eight cents each. Thus, twenty-seven dollars and sixty cents was expended for the mailing of questionnaires.

As far as manhour expenditures were concerned, 167 telephone calls were placed to encourage response. This procedure was employed as an alternative to repeated mailings which require longer time frames and for which there is documented research demonstrating limited success. The average time required per call--this included clarifying telephone numbers, busy signals, no responses, etc.--as recorded on data sheets was approximately 12 minutes. Based on this finding, 33 hours and 24 minutes of clerical services were expended by research personnel on follow-up of questionnaires. This means that 24 minutes of time were required to obtain a completed questionnaire. This finding should dispel any belief that fewer human resources are required in a mailed questionnaire unless one considers repeated mailings. Manhour expenditures not previously mentioned included labeling and stuffing envelopes, checking returned questionnaires, and assimilating information obtained through telephone follow-up data. These totaled approximately 12 ½ clerical manhours.

The telephone survey, as conducted in this study, required no immediate out-of-pocket outlay. The manpower expenditure was very involved, hence

separate time analyses were prepared for non-respondents and respondents before combining the two analyses to obtain the total cost for this treatment.

Three hundred and three calls were made in obtaining the responses of 108 former students. The actual interviews taken over the telephone required approximately 15 minutes each for a total of 27 hours of research assistant time. The preliminary calls to reach the 108 respondents (65.9 percent response rate) averaged four minutes each for a total of 13 hours of research assistant time. Hence, the actual average number of minutes required to obtain a response from the group of respondents was 22 minutes.

One hundred and sixty-four telephone calls were placed to non-respondents and in only 27 cases was it found to be impossible to reach the subject via telephone. The average time spent on each call, as recorded on cover sheets, was six minutes. Therefore, a total of 15 hours and 12 minutes of research assistant time was expended on attempts to obtain an interview from 56 former students. Combining the above manpower data shows that 55 hours and 12 minutes at five dollars per hour were devoted to this method of obtaining follow-up data for a total cost of two-hundred seventy-six dollars.

The personal interview technique was similar to the telephone approach in that there existed no out-of-pocket costs to the investigator. However, for purposes of making comparisons, interviewer mileage was converted to dollar figures at a rate of ten cents per mile. In addition, there were several other factors such as location and time of interview, preparation for interview, and clerical organization that required special attention.

Sixty per cent of all interviews were conducted at the Skyline Center during normal (8:00 A.M. to 5:00 P.M.) working hours. Twenty-eight interviews were completed at the respondents' homes at varying times with the majority

occurring in the evening. The average distance from Skyline was seven miles one-way. However, by scheduling and obtaining interviews by area, mileage was kept to a minimum. Thus, 284 miles were logged in conducting interviews at respondents' homes. Three additional trips, totaling 18 miles, were made to obtain interviews at the subjects' place of employment. Monetary costs for travel, on the basis of ten cents per mile, came to thirty dollars and twenty cents. Travel time for interviewers totaled 12 ½ hours. It is also noteworthy that while appointments made with some individuals for interviews at Skyline were missed, no trips to homes or places of employment proved fruitless.

The average time required for the interview itself was 20 minutes. However, 495 telephone calls were placed, 93 appointments were made, and approximately 12 hours of clerical time was spent in planning and assimilating various pieces of information related to the personal interviews.

Telephone calls averaged eight minutes each, including time spent in looking up changed numbers, etc. Therefore, a total of 116 ½ manhours, 38 ½ by research assistants and 78 by clerical personnel, was required for the personal interviews at a cost of four-hundred twenty-six dollars and fifty cents.

Disregarding those costs that applied equally to all three treatments, it was found that (1) the mailed questionnaire required approximately 46 manhours and twenty-seven dollars and sixty cents, for a total cost of one-hundred sixty-five dollars and sixty cents; (2) the telephone survey required approximately 55 manhours and no out-of-pocket expenditures (this is, of course, subject to local telephone rates and the availability of telephones for interviewing purposes), totaling two-hundred seventy-six

dollars; and (3) the personal interview method required approximately 116 manhours and thirty dollars and twenty cents for a total cost of four-hundred fifty-six dollars and seventy cents.

Question Four

The fourth question investigated was: Which method or combintaion of methods was most cost-effective for conducting a large-scale follow-up study?

There are two ways of examining cost-effectiveness. The more simple approach is to simply compare rate of response with total cost on an individual treatment basis. A more complex method of assessing cost-effectiveness is to weigh the nature or quality of response in conjunction with rate of response and arbitrarily determine this value against cost.

When comparing cost with rate of response, the following ranking of cost-effectiveness was obtained:

1. The mailed questionnaire with subsequent follow-up required 46 manhours and twenty-seven dollars and sixty cents, at a total cost of one-hundred sixty-five dollars and sixty cents, to obtain 83 responses out of a possible 164. This meant that on a per response basis, the cost was approximately two dollars. Had no follow-up been employed, the cost would have been .20 manhours and forty-two cents per response. Thus, the response rate would have been only 38 per cent but the cost-effectiveness was considerably higher.

2. The telephone survey had the highest rate of response, 108 out of 164, and a total cost of two-hundred and seventy-six dollars. The cost on a per response basis was two dollars and fifty-five cents with no other exclusive treatment costs.

3. The personal interview approach had a combined cost of four-hundred fifty-six dollars and seventy cents, including manpower and out-of-pocket costs, to obtain responses from 78 of 164 individuals. On a per response basis, this meant that five dollars and eighty-six cents was expended. Hence, this method was obviously the least cost-effective when one examines only rate of response.

The inclusion of nature or quality of response into a comparison of cost-effectiveness requires that value judgments be made regarding lack of response, extent of response, and willingness of the respondent to express criticisms. Each of these issues must be weighed in such a way that inferences regarding effectiveness may be drawn. It was the opinion of the researcher that obtaining a response and the subsequent validity of that response were most important. This latter concern dealt with the respondents' willingness to express criticism. The extent or dimension of the response (such as number of words or thoughts) was considered to be of relatively less importance.

Based on the above considerations, in conjunction with rate of response, the following interpretation of cost-effectiveness appears warranted. The telephone survey consistently had substantially fewer non-responses than the mailed questionnaire with no lack of willingness to express criticism when responding; hence, the telephone survey would be considered the most cost-effective particularly when one also considers the difference in rate of response as compared to the mailed questionnaire. The low response rate for the personal interview, combined with noticeably fewer critical comments indicated that this method was the least cost-effective. The arbitrary nature of adjusting on the basis of one's personal interpretation of the importance and weighting of quantity versus quality of responses may allow

other rankings of cost-effectiveness; however, the above interpretation appeared quite reasonable on the basis of the data obtained.

Question Five

The fifth question addressed was: Were some methods more effective than others for certain ethnic groups?

It was expected that the three treatments would elicit significantly different response rates from respondents who were members of minority groups. There did exist a statistically significant difference among the treatments when considering all students. Based on this finding, comparisons were made of the three treatments for each ethnic group.

A statistically significant difference was found among ethnic groups. Additional examination revealed that the difference could be attributed to the higher response rate obtained from Anglos.

Comparisons by pairs of ethnic groups showed no difference between Blacks and Mexican-Americans when consolidating respondents across all treatments.

Findings concerning rate of response by ethnicity showed that all methods of survey research employed in the study generally were more effective with Anglos than with Blacks or Mexican-Americans. However, the response from Anglos to the telephone survey method was statistically significant ($\chi^2 = 8.64$, $df = 1$, $p < .005$) when compared to the personal interview approach and was a contributing factor to the differences identified.

While no significant differences were obtained in the examination of treatment comparisons for Blacks and Mexican-Americans, the small sample size, particularly for Mexican-Americans, made it difficult to obtain statistically significant differences. The observable differences may only

be chance occurrences, but the response to telephone contact was very encouraging, especially when one considers the poor response rate to the mailed questionnaire.

It was apparent that the telephone survey elicited a higher response rate for Anglos and, while it was not significantly better for Blacks and Mexican-Americans from a statistical standpoint, the direction of differences favored the telephone survey.

Discussion and Interpretation

The conclusions discussed in this section were based upon findings resulting from the questions posed and must take into account the limitations and procedural considerations of the study.

No Substantive Differences Based on Treatment

Of the 48 subjects randomly selected to receive a second treatment, 19 did in fact respond to two methods of the follow-up survey. It is noteworthy that as with the study in general the telephone interview technique had the greatest number of respondents for the second treatment. It was also found that four subjects who had failed to respond to the initial treatment did respond to the second treatment. This included three individuals who could not be reached by telephone or in person due to either location or lack of telephone service that returned the mailed question, and one former student who had failed to return the mailed questionnaire but was interviewed over the telephone.

As far as substantive differences due to the treatment received, there was no consistency in the responses obtained to support any arguement that the same individual would respond differently to the same question depending

on method employed to obtain said response. Certainly the small sample size makes it difficult to unequivocally state that no difference should occur; however, no indication of differences was realized.

Differences in the Level of Item Non-response across Treatments

There was no support for the Hochstim and Athanasopoulos (1970) finding concerning the completeness of questionnaires being equal across all treatments. There was a statistically significant decrease in the number of subjects choosing not to respond to various questions when personalized methods of follow-up were administered. It is probable that the subjects of the earlier study were not representative of former high school students from a large urban school district such as Dallas.

Comparability of Findings across Treatments

Consistent with the findings of Hochstim and Athanasopoulos (1970) but in opposition to those of Jackson and Rothney (1961) results of the present study suggest that once responses are obtained via any of the strategies there is no reason to expect more complete or illuminating responses because of personal contact or anonymity of respondents. Close-ended questions were found to elicit almost identical responses except in the case of two ratings involving the quality of instruction and counseling. These instances demonstrated that the mailed questionnaire does obtain a greater proportion of negative responses than at least one of the personalized methods; however, there was no consistency in the findings of the present study to support the mailed questionnaire over either the telephone or personal interview techniques. On open-ended questions for which responses were obtained, the results demonstrate that there is no reason to expect appreciably different responses for any of the treatments.

Differences in Rate of Response across Treatments

Related studies had produced findings which indicated that the rate of response should not differ significantly for any of the follow-up strategies employed. The present study found this to be true for the mailed questionnaire and personal interview; however, the telephone interview technique did obtain a significantly higher response rate particularly for Anglos when compared with the personal interview. While statistical significance was not found in the differences among treatments for Blacks and Mexican-Americans, this may be attributed to the small number of minority subjects who actually responded. The lack of minority representation with regard to interviewers (both telephone and personal) may have also contributed to this lack of statistical significance when comparing personalized methods with the mailed questionnaire. Across all ethnic groups the results of the present study support increased awareness of the value of telephone interviewing for achieving higher response rates.

Differences in Cost across Treatments

The cost analysis findings of the present study are consistent with previously reported research (Hochstim and Athanasopoulos, 1970; Willardsen, 1972; Orr and Neyman, 1965; and Jackson and Rothney, 1961). The personal interview technique is over twice as expensive as either the mailed questionnaire or telephone interview. The costs for the latter two were found to be quite similar when compared on the basis on the number of responses obtained by each. The other cost aspect that was important to understand is that much of the cost incurred in follow-up research involves human resources. This is particularly important when one considers the type of individuals required

for various tasks. The mailed questionnaire except for analysis can be conducted almost solely by clerical personnel once instrumentation has been developed. Similarly telephone interviewing can be undertaken by clerical personnel who have been given training as to appropriate procedures. Personal interviewing does seem to require specialized skills that are most frequently found in college graduates.

Therefore cost analyses are deceiving unless consideration is paid to what kind of costs are being incurred such as capital outlay versus human costs. This is combined with the above mentioned variety of costs tied to human resources which can drastically affect overall costs of the survey.

Implications for Future Follow-up Surveys

Considering the need for constructively designed follow-up research to provide information for improving existing programs and formulating new programs, the ultimate goal of the present study was to make meaningful suggestions for insuring the highest quality of resultant data.

For large school districts and programs with adequate staffing patterns, the conduct of a follow-up survey may have greater flexibility than small districts whose need for such evaluation is no less. The results of the present study supported by related research indicates that more attention should be paid to telephone interviewing, a technique suitable for both large and small districts. The limitations involving manpower requirements and locality of the potential respondents can be counteracted by incorporating mailed questionnaires into a study for those subjects unavailable by telephone.

Procedures to incorporate in any follow-up study to insure the best possible results should include the following:

1. Careful identification of the population to be studied paying particular attention to obtain accurate addresses and telephone numbers. Frequently follow-up studies are not adequately planned so that there is difficulty in locating former students.

2. Greater use of clerical personnel whenever and wherever feasible without sacrificing quality. This would require taking necessary steps to provide training and leadership for such individuals to insure the quality of their work. There is no need to use high salaried individuals for tasks perfectly suitable for more cost-effective personnel.

3. Instrumentation that is goal oriented so that results of the study may be tied to program improvement policies. Traditionally follow-up studies only seek answers to simplistic questions having little potential influence for school improvement regardless of the findings. Once subjects are contacted, it costs no more to ask in depth evaluative questions.

In conclusion, the study conducted here indicates that a carefully conducted telephone interview survey performed largely by clerical personnel combined with mailed questionnaires to subjects living outside the immediate community or those not having telephones should insure a response rate of at least 70 per cent. Furthermore, the nature of the information obtained via this process should be equivalent to the much costlier data achievable from a personal interview approach.

FOOTNOTES

[1]Robert M. Jackson and J.W.M. Rothney, "A comparative Study of the Mail Questionnaire," Personnel and Guidance Journal, 39, (March, 1961), p. 569.

[2]Miriam B. Gurr, "A Continuing Follow-up Study," Journal of Secondary Education, 26 (May, 1951), pp. 302-305.

[3]Claude W. Grant, "A Follow-up Study of Spring, 1966, High School Graduates in the State of Utah," Personnel and Guidance Journal, 47 (October, 1968) pp. 157-162.

[4]Arthur L. Maser, "Highline Graduates Survey - 1970," unpublished report, Seattle, Washington, (March, 1971).

[5]Paul J. Casawant and Sandra L. Rudin, "Summary: Senior Questionnaire Results, June, 1970," unpublished report, Syracuse, New York, (October, 1970).

[6]"Education for the 70's," unpublished research report, Tacoma Public Schools, Tacoma, Washington, (May, 1971).

[7]Frederick G. Nichols, "Criticism, Comment, and Challenge," The Journal of Business Education, 18 (October, 1942), p. 9.